Validation of Handheld Ultrasound Devices for Point of Care Use in Rheumatology Study

GRAPPA Ultrasound Working Group

Statistical Analysis Plan

Principal Investigators

Sibel Zehra Aydin, Lihi Eder, Gurjit Kaeley

Statistical Analysis Plan Author

Seyyid Bilal Acikgoz

11/05/2023

Study Protocol Version 2.0 (21/03/2023)

Table of Contents

1. **Introduction** 3
   - **1.1. Background and Rationale** 3
   - **1.2. Objectives** 4
2. **Study Methods** 6
   - **2.1. Trial Design** 6
   - **2.2. Randomization and Blinding** 6
   - **2.3. Sample Size** 7
   - **2.4. Statistical Interim Analyses and Stopping Guidance** 7
   - **2.5. Timing of Final Analysis** 7
   - **2.6. Timing of Outcome Assessment** 7
3. **Statistical Principles** 8
   - **3.1. Confidence Intervals and P-values** 8
   - **3.2. Adherence and Protocol Compliance** 8
   - **3.3. Analysis Populations** 8
4. **Trial Population** 8
   - **4.1. Eligibility** 8
   - **4.2. Withdrawal/Follow-up** 9
   - **4.3. Baseline Patient Characteristics** 9
5. **Analysis** 10
   - **5.1. Outcome Definition** 10
     - **5.1.1. Primary Outcome** 10
     - **5.1.2. Secondary Outcomes** 11
       - **5.1.2.1. Joints** 11
       - **5.1.2.2. Nails** 12
       - **5.1.2.3. Tendons** 12
       - **5.1.2.4. Entheses** 13
     - **5.1.3. Safety Outcomes** 16
   - **5.2. Analysis Methods** 16
   - **5.3. Missing Data** 16
6. **References** 17

1. Introduction

1.1. Background and Rationale

Recently, ultrasonography (US) has experienced a rapid evolution in the field of rheumatology, an evolution certainly driven by the broad applicability of point-of-care (POC) US for the assessment of rheumatic diseases. Despite the use of US in many other medical fields, such as gynecology, emergency or gastroenterology for much longer, the uptake by the rheumatology community was only initiated in the last two decades. This is mostly due to the fact that most of the structures that are assessed in rheumatology are very superficial, and the technology to generate a high resolution view for these superficial structures, to a degree of being able to detect even mild inflammatory changes, have been developed relatively recently.

Nowadays, rheumatologists typically use US for guided injections and for the assessment of joint structures, connective and vascular tissues, and related pathologies.[1] When it comes to the assessment of musculoskeletal (MSK) structures, the value of US lies upon a unifying principle for many arthritides: rapid detection of highly relevant and often times subclinical features of disease. For example, in a rapidly progressing disease like Rheumatoid Arthritis (RA), US allows faster detection of synovitis and bone erosion, resulting in earlier fulfillment of diagnostic criteria.[2]

Early disease detection with US perhaps takes its most relevant sense in psoriatic arthritis (PsA), where efforts are currently being deployed to intercept disease in its transition from psoriasis (PsO).

Moving on to the subject of ultrasonography scanners as medical devices; more specifically the high quality US scanners and high-resolution transducers required for the practice of rheumatology. A number of barriers persist and stand in the way of a wider use of POC US for the detection, diagnosis and management of rheumatic diseases. A significant one is the acquisition and maintenance costs of high quality instruments. There are recent technical and technological advances in the field of handheld ultrasonography that are set to overcome the access barrier. Typically, the cost of acquisition of scanners with a greyscale frequency of at least 13 MHz and a Doppler frequency of at least 8 MHz ranges from 25000CAD to

90000CAD. Hand-held US technology promises to take this cost down to a price tag under 10,000CAD with the introduction of affordable high-definition scanners possessing the specifications/requirements for use in the rheumatology practice (Greyscale frequency: 12 - 20 MHz and Doppler frequency: 8 - 12 MHz). Clarius Mobile Health Inc., an innovative company based out of Vancouver B.C., produces such devices. Clarius US scanner have regulatory approval by the FDA and Health Canada. However, before they can be specifically used for the practice of rheumatology, their performance needs to be validated against gold-standard devices for key interventions.

US holds significant promise as an imaging tool in rheumatology and in the future, handheld US could be used at the bedside to provide diagnostic and prognostic information, as well as guide both systemic and local treatment decisions and applications. Through the accurate assessment of disease extension and activity (e.g. presence of enthesitis in a PsA patient presenting with primarily as synovial disease), the bedside application of US in standard assessment of the PsA patients will enable the understanding of which domains are involved and would require treatment/and which treatment based on the domain involved. There are numerous advantages of the hand-held US devices over the existing gold standard devices for being accessible by more physicians, therefore by more patients. The ability to carry the device in their pocket will allow the physicians to be able to use in multiple settings, e.g. in different offices or inpatient vs outpatient clinics. Making the bedside US a part of the clinical assessment will avoid any delays in diagnosis and lead to earlier treatments. In addition, it increases the patients adherence to therapy adjustments.[3] As such, the goal of this trial will be to validate two affordable handheld MSK US scanner against gold-standard devices through an assessment of their accuracy for:

- visualizing anatomical structures and pathologies
- detecting vascular flow

1.2. Objectives

Our aim to test the concurrent validity of the Clarius handheld US devices versus gold-standard device to detect characteristic features of healthy and rheumatic joints (i.e. anatomical structures and vascular flow.

Primary objective:

To determine whether handheld US devices (Clarius HD3 L20 and L15 scanners) B mode are as accurate as gold standard device (GE Logic E9/S8) at visualizing intraarticular synovitis

Secondary objectives:

1. To determine whether handheld US devices (Clarius HD3 L20 and L15 scanners) power Doppler mode are as accurate as gold standard device (GE Logic E9/S8) at detecting intrasynovial signals

2. To determine whether handheld US devices (Clarius HD3 L20 and L15 scanners) B mode are as accurate as gold standard device (GE Logic E9/S8) at visualizing tenosynovitis

3. To determine whether handheld US devices (Clarius HD3 L20 and L15 scanners) power Doppler mode are as accurate as gold standard device (GE Logic E9/S8) at detecting intratendineous signals

4. To determine whether handheld US devices (Clarius HD3 L20 and L15 scanners) B mode are as accurate as gold standard device (GE Logic E9/S8) at visualizing bone erosions

5. To determine whether handheld US devices (Clarius HD3 L20 and L15 scanners) B mode are as accurate as gold standard device (GE Logic E9/S8) at grading intraarticular synovitis

6. To determine whether handheld US devices (Clarius HD3 L20 and L15 scanners) power Doppler mode are as accurate as gold standard device (GE Logic E9/S8) at grading intrasynovial signals

7. To determine whether handheld US devices (Clarius HD3 L20 and L15 scanners) B mode are as accurate as gold standard device (GE Logic E9/S8) at grading bone erosions

8. To determine whether handheld US devices (Clarius HD3 L20 and L15 scanners) power Doppler mode are as accurate as gold standard device (GE Logic E9/S8) at visualizing elementary lesions of enthesitis

9. To determine whether handheld US devices (Clarius HD3 L20 and L15 scanners) power Doppler mode are as accurate as gold standard device (GE Logic E9/S8) at detecting entheseal signals

10. To determine whether handheld US devices (Clarius HD3 L20 and L15 scanners) B-mode are as accurate as gold standard device (GE Logic E9/S8) at detecting trilaminar appearance of the nail plate abnormalities

2. Study Methods

2.1. Trial Design

Multicenter and single arm study (repeating an image collection protocol with hand held US and the gold standard devices for each patient)

2.2. Randomization and Blinding

Study patients will not be randomized to any group. The same anatomical regions will be evaluated with the same ultrasound protocol for each patient included in the study. The scoring of the US images will be done blindly by the principal investigator at OHRI as stated below.

At the central site (OHRI), the research assistant will give a unique identifier number to each image, for a random quality control and for cross referencing whenever needed. The cropped images, as detailed below, will not have the subject ID visible to the PI at the time of reading but will be accessible for the quality control. (read-only access). The research assistant at OHRI is the only site personnel who has the capacity to uncrop the images in the PowerPoint file (password protected files).

Images will not contain any identifiable information such as Date of Birth (DOB) or initials. The US images will be transferred to a PowerPoint file by the research assistant at OHRI in JPEG format. The research assistant will generate an unblinded master list, inaccessible to other site personnel, to link the slide numbers with the patients and scanned anatomical sites and the slide will have no other information on the patient number or ID. For scoring the images by the PI, a random order slide show will be conducted, irrespective of the machine used or the anatomical site or patient assessed, to ensure blindness to data related to the

patient identifiers (The PI will not be blinded to the machine that the image was taken with as the JPEG format that is achieved from different machines are identifiable, but due to the random order scoring, images that belong to the same joint by the different machines are not to be scored consecutively). There will be nine separate powerpoint files, for images of joints, tendons, entheses, nail including power Doppler and gray scale findings; and grey scale file for erosions.

2.3. Sample Size

Study agreement analyses will be done per joint. To give a kappa ≥ 0.61 (from substantial interrater agreement to almost perfect agreement) with a confidence interval width of 0.15 and the expectation of approximately 20 % of joints assessed having any synovitis in B mode (based on a previous study by our group), 683 joints would be required.[4] This corresponds to 30 patients if 24 joints per patient are assessed. Non-inferiority margin will be $\kappa \geq 0.61$.

2.4. Statistical Interim analyses and stopping guidance

An interim analysis will be done after 10 patients. In the interim analysis, the agreement between the images obtained with the handheld ultrasound and the images on the gold standard device will be compared after the blinded scoring. A moderate level of agreement (defined as a kappa value of >0.40) between the two tools in the interim analysis will allow parallel initiation of subsequent studies proposed in the research framework. The results of this interim analysis will not be used to adjust the design of the remainder of the study. We do not have any formal stopping rules for the trial.

2.5. Timing of final analysis

After all sites recruit 10 patients each, the data will be verified and locked. Final analysis will commence once the final lock has been confirmed by the Principal Investigator.

2.6. Timing of outcome assessment

First outcome assessment will be made when 10 patients are included in the study for interim analysis. The final outcome assessment will be made when the planned total number of patients is reached.

3. Statistical Principles

3.1. Confidence Intervals and P-values

We will consider $p< 0.05$ statistically significant for our outcomes. Results will be presented with their values with the 95% confidence interval.

3.2. Adherence and protocol compliance

Attention will be paid to compliance with the protocol at every stage of the study. 100% compliance with the protocol is required. All data will be evaluated twice for protocol compliance. US images and CRFs uploaded from each center will be reviewed by the research assistant at Ottawa Hospital Research Institute (OHRI) center. If there are any missing or erroneous data in the CRF copy, research assistant will contact the site to ensured that the errors and deficiencies are corrected in source document according to Good Documentation Practice. Site should then re-scan the CRF to the SharePoint, ensuring sequential versioning. Any queries will be confirmed with the site within a week of the data entry. Then the analyzed CRFs data will be transferred to Research Electronic Data Capture (REDCap) (version 12.4.18 - © 2023 Vanderbilt University) by the research assistant at OHRI. Before each analysis (interim and final), all the paper CRFs and REDCap data will be compared for quality assurance. Twenty percent of data monitoring will be performed, as per monitoring plan. If a protocol violation is detected in obtaining any data, this data will documented and retained in records.

3.3. Analysis populations

All patients will be included in the analysis.

4. Trial Population

4.1. Eligibility

Adult patients with peripheral PsA who are not in minimal disease activity (MDA) and show at least one tender and swollen joint are eligible for study.

Inclusion Criteria

- Age ≥18
- Meets the classification for psoriatic arthritis (CASPAR) criteria
- Able to provide an informed consent
- Having peripheral disease phenotype of PsA
- At least one tender and swollen join on the day of US

Exclusion Criteria

- Having isolated axial PsA
- Being in MDA with no tender and swollen joints

4.2. Withdrawal/Follow-up

There are no follow-up visits in the study, therefore we do not expect a withdrawal.

4.3. Baseline Patient Characteristics

There will be only one study visit and the following data will be collected at the baseline visit.

- Patient's age
- Gender
- Body Mass Index (BMI)
- Physical activity
- Smoking
- Treatments and disease activity
- 66/68 joint count for tender and swollen joints
- Spondyloarthritis Research Consortium of Canada (SPARCC) enthesitis index,
- Number of dactylitic digits,
- Body surface area for skin disease
- Presence of nail disease
- Pain Visual Analogy Scale (VAS),

- Patient global assessment
- Health Assessment Questionnaire Disability Index (HAQ-DI)
- Erythrocyte sedimentation rate (ESR)
- C-Reactive Protein (CRP)
- Adverse events related to ultrasound scanning

5. Analysis

5.1. Outcome Definition

5.1.1. Primary Outcome

After the imaging of all patients is completed, intraarticular synovitis scoring will be done based on the OMERACT definitions (scales of 0-3) as follows, for the anatomical sites given in Table 1.[5,6]

Greyscale inflammatory (hypoechoic) synovial hyperplasia:

Grade 0: no hypoechoic synovial hyperplasia

Grade 1: minimal hypoechoic synovial hyperplasia (filling the angle between the periarticular bones, without bulging over the line linking tops of the bones)

Grade 2: hypoechoic synovial hyperplasia bulging over the line linking tops of the periarticular bones but without extension along the bone diaphysis

Grade 3: hypoechoic synovial hyperplasia bulging over the line linking tops of the periarticular bones and with extension to at least one of the bone diaphysis

Table-1: Probes/scanners, views and anatomical sites for primary outcomes

Anatomical site	Views	GE Logic E9/S8	Clarius L15	Clarius L20
2nd MCP joints	dorsal and lateral	X	X	X
3rd MCP joints	dorsal	X		X
2nd-3rd PIP, 2nd-3rd DIP	dorsal	X		X
Wrist	dorsal	X	X	X
5th MTP joints	dorsal and lateral	X		X
Elbow joints	posterior	X	X	
Shoulder joints	posterior	X	X	
Knee joints	anterosuperior	X	X	
Ankle joints	anterior	X	X	

5.1.2. Secondary Outcomes

5.1.2.1. Joints

After the imaging of all patients is completed, the intraarticular Doppler scoring will be done based on the OMERACT definitions (scales of 0-3) (Table 1):[5,6]

Power Doppler signal:

Grade 0: no flow in the hypoechoic synovial hyperplasia

Grade 1: up to three single spots signals or up to two confluent spots or one confluent spot plus up to two single spots

Grade 2: vessel signals in less than half of the area of the synovium (≤50%)

Grade 3: vessel signals in more than half of the area of the synovium (>50%)

Also, a lateral view of the 2nd metacarpophalangeal and 5th metatarsophalangeal joints will be recorded in the baseline visit for erosion assessment. Erosions will be described as intra-articular discontinuity of the bony surface seen in 2 perpendicular planes and will be evaluated after the imaging of all patients is completed.

5.1.2.2. Nails

B mode and doppler images of the 2nd nail or most involved nail with all probes (Clarius L20, GE Logic E9/S8) will be recorded in the baseline visit. The loss of the trilaminar appearance and the presence of doppler signal in the nail will be evaluated after the imaging of all patients is completed. During imaging of the nail, if any Doppler signal is detected in nail bed, regardless of its severity, the presence of Doppler signal will be considered positive.

5.1.2.3. Tendons

B mode and Doppler images of the tibialis posterior tendon with Clarius L15 and GE Logic E9/S8 will be recorded in the baseline visit. The presence of tendonitis, tenosynovitis and intratendinous Doppler signal in tibialis posterior tendon will be evaluated after the imaging of all patients is completed. Tendonitis will be defined as loss of fibrillary echotexture with hypoechogenicity within the tendon fibres on B-mode (regardless of the presence of Doppler signal). Tenosynovitis will be defined as characterized by hypoechoic or anechoic thickened tissue with or without fluid in the tendon sheath on B-mode (regardless of the presence of Doppler signal). If any Doppler signal is detected in tendon, regardless of its severity, the presence of intratendinous Doppler will be considered positive.

B mode and Doppler images of the 2nd extensor digitorum tendon with Clarius L20 and GE Logic E9/S8 will be recorded in the baseline visit. The presence of paratenonitis and intratendinous Doppler signal in the extensor digitorum tendon will be evaluated after the imaging of all patients is completed. Paratenonitis will be defined as the lack of a sheath on the extensor tendon above the metacarpophalangeal joint with accompanying inflammatory changes to the extensor tendon consisting of increased thickness, loss of fibrillary echotexture, with or without power Doppler signal. If any Doppler signal is detected in the tendon or the tendon sheath, regardless of its severity, the presence of intratendinous Doppler will be considered positive.

5.1.2.4. Entheses

After the imaging of all patients is completed, Elementary lesions of enthesitis will be defined and scored as per the GRAPPA US working group's definitions, as used in the multicenter DUET study (Table 2):[4]

All of the elementary lesions will be assessed for their presence or absence. In addition, some of the lesions will also be scored for their severity using a semi-quantitative system (grade 0 to 3). The definitions of each grade are outlined below.

Table-2: Probes/scanners, views and anatomical sites for enthuses assessment

Anatomical site	**Views**	**GE Logic E9/S8**	**Clarius L15**	**Clarius L20**
Achilles tendon insertions	posterior	X	X	X
Supraspinatus tendon insertions	lateral	X	X	
Triceps tendon insertions	posterior	X	X	
Common extensor tendon origins	lateral	X	X	
Quadriceps tendon insertions	anterior	X	X	
Patellar ligament (origin and insertion)	anterior	X	X	
Plantar fascia insertions	plantar	X	X	
Extensor digitorum tendon insertions at DIP and PIP	dorsal	X		X

The following sonographic entheseal lesions should be scored:

- **Hypoechogenicity**: Distinct loss of homogenous fibrillar pattern with relative hypoechogenicity compared to the rest of the enthesis after correcting for anisotropy.

  Grade 0: Absent

  Grade 1: Present

- **Thickening:** Increased thickness of the tendon/ligament at the enthesis compared to its body. Thickness may be difficult to judge and should be suspected when accompanied by other entheseal lesions.

  Grade 0: Absent

  Grade 1: Present

- **Bone Erosion:** A cortical defect confirmed with a step-down contour defect detected in two planes at the insertion of the tendon/ligament to the bone.

  Grade 0: Absent

  Grade 1: Present

- **Enthesophyte:** A step-up bony prominence at the normal bone contour. Grade the severity of the enthesophyte based on its length. Although the readers are not expected to measure the length of the enthesis we provided suggested cut off points to guide the grading of the enthesophytes

  Grade 0: No enthesophytes

  Grade 1: Small enthesophyte

  Grade 2: Medium enthesophyte

  Grade 3: Large enthesophytes

- **Calcification:** Hyperechoic linear structures detected within the tendon/ligament at the insertion to the bone but with no congruency with the bone.

  Grade 0: No calcifications

  Grade 1: Punctate hyperechoic area

  Grade 2: Linear calcification without acoustic shadow

  Grade 3: Egg-shell calcification with posterior acoustic shadow

- **Doppler Signal:** The presence of positive Doppler signal at the enthesis, confirmed in two perpendicular planes and distinguished from reflection of surface artifacts and nutritional vessel signal.

  The intensity of the Doppler signal at the enthesis will be graded using a semi-quantitative score.

  Note: We will consider any Doppler signal at the enthesis area including signals appearing beyond 2 mm of the bony cortex.

  We will score the intensity of the Doppler signal on a semi quantitative grading system:

  Grade 0: No Doppler signal

  Grade 1:A single confluent Doppler signal or up to 3 discrete Doppler spots

  Grade 2: Doppler signal affecting less than half of the enthesis

  Grade 3: Doppler signal covering more than half of the enthesis

  In addition, the location of each Doppler signal will be recorded:

  Zone 1: ≤2mm from the bone cortex

  Zone 2:>2mm from the bone cortex

5.1.3. Safety Outcomes

Adverse events will be recorded during scanning.

5.2. Analysis Methods

Demographic, clinical and disease characteristics of patients will be presented using descriptive statistics. Numeric data showing normal distribution will be shared with mean and standard deviation, and those without normal distribution will be shared with median and interquartile range. Categorical data will be shared with numbers and percentages.

The primary endpoint analysis will be the interrater agreement of detecting any synovitis in B mode with the Clarius and gold standard machine. The kappa coefficients will be evaluated using the guideline outlined by Landis and Koch, where the strength of the kappa coefficients are: 0.01-0.20 slight; 0.21-0.40 fair; 0.41-0.60 moderate; 0.61-0.80 substantial; 0.81-1.00 almost perfect.[7]

For secondary outcomes, the interrater agreement for the presence of Doppler signals within the joints, tenosynovitis, erosions, nail, as well as features of enthesitis (hypoechogenicity, thickening, erosions, enthesophytes, calcifications) will also be evaluated using the same method. The agreement of the semiquantitative grading of the intraarticular findings' severity (synovitis in B mode, Doppler signals, erosions, each being on a scale between 0-3) will be done using weighted kappa analysis.[8]

5.3. Missing Data

If there are missing images for some sites for any of the probes, the images that were obtained for the same site using other probe(s) will also be excluded. The number of missing joint/tendon/entheses and nail images will be reported. Missing data will not be imputed.

6. References

1-Kang T, Emery P, Wakefield RJ. A brief history of ultrasound in rheumatology: where we are now. Clin Exp Rheumatol. 2014 Jan-Feb;32(1 Suppl 80):S7-11. Epub 2014 Feb 17. PMID: 24529159

2-Filer A, de Pablo P, Allen G, Nightingale P, Jordan A, Jobanputra P, Bowman S, Buckley CD, Raza K. Utility of ultrasound joint counts in the prediction of rheumatoid arthritis in patients with very early synovitis. Ann Rheum Dis. 2011 Mar;70(3):500-7. doi: 10.1136/ard.2010.131573. Epub 2010 Nov 29. PMID: 21115552; PMCID: PMC3033529.

3-Tan YK, Teo P, Saffari SE, Xin X, Chakraborty B, Ng CT, Thumboo J. A musculoskeletal ultrasound program as an intervention to improve disease modifying anti-rheumatic drugs adherence in rheumatoid arthritis: a randomized controlled trial. Scand J Rheumatol. 2022 Jan;51(1):1-9. doi: 10.1080/03009742.2021.1901416. Epub 2021 Jun 10. PMID: 34107851.

4-Tom S, Zhong Y, Aydin SZ, Development of a Preliminary Ultrasonographic Enthesitis Score in Psoriatic Arthritis - GRAPPA Ultrasound Working GroupJRheumatol, 2019 Apr;46(4):384-390

5-D'Agostino MA, Boers M, Wakefield RJ, et. al., Exploring a new ultrasound score as a clinical predictive tool in patients with rheumatoid arthritis starting abatacept: results from the APPRAISE study, RMD Open. 2016; 2(1), 1-7

6-D'Agostino MA, Wakefield RJ, Berner-Hammer H, et. al., Value of ultrasonography as a marker of early response to abatacept in patients with rheumatoid arthritis and an Validation of Handheld Ultrasound Devices for Point of Care Use in Rheumatology, Study Protocol Version 2.0 March 21,2023 24 / 24 inadequate response to methotrexate: results from the APPRAISE study, Ann Rheum Dis. 2016 Oct; 75(10): 1763–1769

7-Landis JR, Koch GG, The measurement of observer agreement for categorical data, Biometrics, 1977, 33(1):159-74

8-Cohen J, Weighed kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. Psychological Bulletin.1968, 70 (4): 213–220